

**Regionalization of
patterns of flow
intermittence**

T. H. Snelder et al.

Regionalization of patterns of flow intermittence from gauging station records

**T. H. Snelder^{1,2}, T. Datry¹, N. Lamouroux¹, S. T. Larned³, E. Sauquet⁴, H. Pella¹,
and C. Catalogne⁴**

¹IRSTEA, UR MALY, Villeurbanne, France

²Aqualinc Research, Christchurch, New Zealand

³National Institute of Water and Atmospheric Research, Christchurch, New Zealand

⁴IRSTEA, UR HHLY, Villeurbanne, France

Received: 24 December 2012 – Accepted: 15 January 2013 – Published: 30 January 2013

Correspondence to: T. H. Snelder (t.snelder@aqualinc.co.nz)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Abstract

Understanding large-scale patterns in flow intermittence is important for effective water resource management. We used daily flow records from 628 gauging stations on rivers with minimally modified flows distributed throughout France to predict regional patterns of flow intermittence. For each station we calculated two annual times-series describing flow intermittence; the frequency of zero-flow periods (consecutive days of zero-flow) in each year of record (FREQ; yr^{-1}), and the total number of zero-flow days in each year of record (DUR; days). These time series were used to calculate two indices for each station, the mean annual frequency of zero-flow periods (mFREQ; yr^{-1}), and the mean duration of zero-flow periods (mDUR; days). Approximately 20 % of stations had recorded at least one zero-flow period. Dissimilarities between pairs of gauges calculated from the annual times-series (FREQ and DUR) and geographic distances were weakly correlated, indicating that there was little spatial synchronization of zero-flow. A flow-regime classification for the gauging stations discriminated intermittent and perennial stations, and an intermittence classification grouped intermittent stations into three classes based on the values of mFREQ and mDUR. We used Random Forest (RF) models to relate the flow-regime and intermittence classifications to several environmental characteristics of the gauging station catchments. The RF model of the flow-regime classification had a cross-validated Cohen's kappa of 0.47, indicating fair performance and the intermittence classification had poor performance (cross-validated Cohen's kappa of 0.35). Both classification models identified significant environment-intermittence associations, in particular with regional-scale climate patterns and also catchment area, shape and slope. However, we suggest that the fair-to-poor performance of the classification models is because intermittence is also controlled by processes operating at scales smaller than catchments, such as groundwater-table fluctuations and seepage through permeable channels. We suggest that high spatial heterogeneity in these small-scale processes partly explains the low spatial synchronization of zero-flows. While 20 % of gauges were classified as intermittent, the flow-regime

HESSD

10, 1511–1551, 2013

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



model predicted 39 % of all river segments to be intermittent, indicating that the gauging station network under-represents intermittent river segments in France. Predictions of regional patterns in flow intermittence provide useful information for applications including environmental flow-setting, estimating assimilative capacity for contaminants, designing bio-monitoring programs and making preliminary estimates of the effects of climate change on flow intermittence.

1 Introduction

A large proportion of the river segments on Earth are intermittent, i.e. they periodically cease to flow (Larned et al., 2010b). Most river networks include intermittent segments, which may be concentrated in headwater, mid-catchment, or downstream areas, or interspersed over entire networks (Lake, 2003; Turner and Richter, 2011a). Many river networks in arid regions are entirely intermittent (Jacobson and Jacobson, 2012; Meirovich et al., 1998). Temporal patterns of flow intermittence range from near-perennial flow-regimes with infrequent, short periods of zero-flow to episodic flow-regimes with rare flow events separated by long zero-flow periods (Crocker et al., 2003; Houston, 2006; Larned et al., 2011; Meirovich et al., 1998). In turn, the duration and frequency of zero-flow periods are increasingly viewed as the primary determinants of river ecosystem processes (Corti et al., 2011; Datry et al., 2011; Dieter et al., 2011) and biotic communities (Arscott et al., 2010; Datry, 2012; Davey and Kelly, 2007).

Many intermittent rivers support diverse plant and animal communities, particularly when viewed at time-scales that encompass periods of river flow, standing water, and no water. Aquatic and terrestrial species colonize and emigrate from intermittent rivers in response to shifts between wet and dry habitat (Acuña et al., 2005; Corti and Datry, 2012; Datry et al., 2012; Davey and Kelly, 2007; Steward et al., 2011). A smaller number of intermittent river specialists persist through multiple wet-dry cycles; this group includes aestivating fish and encysting invertebrates (Kikawada et al., 2005; Perry et al., 2008; Sayer, 2005). The alternating occupation of habitat by aquatic and terrestrial

HESSD

10, 1511–1551, 2013

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



HESSD

10, 1511–1551, 2013

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[⏪](#)[⏩](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

species means that the time-averaged biodiversity of intermittent segments can exceed that of perennial segments (Katz et al., 2012). In addition to their ecological values, intermittent rivers provide numerous ecosystem services, including flood irrigation, flood control, and waste-water conveyance (Angel et al., 2010; Larned et al., 2010a). In the Mediterranean region and other water-scarce areas, intermittent rivers represent an important source or the sole source of freshwater for consumptive use (e.g. Jacobson and Jacobson, 2012; Ji et al., 2006). Understanding large-scale patterns in flow intermittence in these regions is a prerequisite for effective water resources management.

The last decade has seen rapid growth in ecological and water resources research focused on intermittent rivers (Datry et al., 2011). This research has been accompanied by a widespread acknowledgment that intermittent rivers require careful management to protect biological and socio-economic values (Jacobson and Jacobson, 2012; Larned et al., 2010a; Nadeau and Rains, 2007). However, current management practices and natural resources policies are often inadequate. For example, there are few restrictions imposed on the use or alteration of intermittent rivers in the United States compared with perennial rivers (Leibowitz et al., 2008). In much of the world, alterations of intermittent river channels and flow-regimes are entirely unrestricted (Elmore and Kaushal, 2008; Gómez et al., 2005; Hughes, 2005). Effective management of intermittent rivers is impeded by the scarcity of information about their abundance, distribution patterns, patterns of flow variability, and the environmental conditions that produce those patterns (Datry et al., 2011; Fritz et al., 2008; Hansen, 2001).

Identifying intermittent rivers and river segments, and characterizing intermittent flow-regimes at scales larger than individual catchments pose several challenges. Small-scale river maps (> 1:25 000) produced from digital elevation models, aerial photography or airborne laser scanning omit many intermittent segments (Hansen, 2001; Leopold, 1994). More intermittent segments are included in detailed maps based on field surveys (e.g. Brooks and Colburn, 2011), but large-scale field-surveys are prohibitively laborious. As alternatives to surveys and remote sensing, empirical models have been used to predict the locations of intermittent river segments based on

catchment characteristics (Bent and Steeves, 2006; Heine et al., 2004). These models perform moderately well in the areas for which they are defined and provide information about intermittence-environment associations over large, environmentally heterogeneous areas. For example, an empirical model used to predict the occurrence of streams identified as intermittent or perennial based on visual observation during low-flow periods in the state of Massachusetts was reasonably accurate, with a misclassification rate of approximately 20 %, (Bent et al., 2006). In another empirical modeling study, intermittent segments in forested headwater streams were distinguished from perennial segments using field survey data (Fritz et al., 2008). However, survey data alone are insufficient for further subdividing intermittent segments on the basis of intermittent flow patterns (i.e. grouping segments with similar frequencies and durations of zero-flow); this step requires flow time-series that field surveys cannot provide.

Grouping gauging stations into flow-regime classes based on time-series data is an important component of large-scale river management and research (Olden et al., 2011). Flow-regime classifications serve as spatial frameworks for environmental monitoring, and they simplify water allocation decisions and environmental flow-setting (Kennard et al., 2010; Olden et al., 2011; Snelder et al., 2009). Most flow-regime classifications are based on hydrological indices calculated from time-series recorded at gauging stations. Hydrological indices describe the magnitude, timing, duration, rate-of-change and frequency of flow events. Statistical similarities in hydrological indices are then used to group gauging stations with similar flow-regimes (Olden et al., 2011). The utility of a flow-regime classification is maximized when class membership can be extrapolated to ungauged locations. In several recent studies, flow time-series have been combined with spatial data describing environmental conditions in statistical classifications that were used to predict river flow-regimes at ungauged locations (e.g. Kennard et al., 2010; Snelder et al., 2009). To our knowledge, there have been no comparable studies that focused specifically on flow intermittence.

In this study we used daily flow records from 628 gauging stations on rivers in France with unmodified or minimally modified flows regimes to study regional patterns of flow

HESSD

10, 1511–1551, 2013

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



intermittence. Our study had two objectives: (1) to characterize flow intermittence in terms of the spatial distribution of intermittent segments, and the frequency and duration of zero-flow periods in those segments; (2) to determine the extent to which flow intermittence patterns are associated with environmental conditions.

2 Materials and methods

2.1 Study area

The study area was continental France, which extends from 50 to 43° N latitude and has an area of 550 000 km². Environmental variation within France that is pertinent to hydrological patterns is summarized by the Hydro-Ecoregions (HER) framework (Wasson et al., 2002). The HER is a regionalization developed for river management in accordance with the European Water Framework Directive. The first level of the HER divides France into 21 regions based primarily on variation in climate, topography and geology. Climate conditions in the hydro-ecoregions range from mediterranean in the “Méditerranéen” and “Cévennes” regions to temperate maritime in the “Armoricain” region, to cold and wet in the alpine “Alpes internes” (Fig. 1). Geological conditions range from calcareous in the “Jura-Préalpes” regions to alluvial in the “Plaine Saône” and “Alsace” regions (Fig. 1). Topographic conditions range from plains in the “Alsace” regions to high mountains in the “Alpes internes” and “Pyrénées” regions.

2.2 Hydrological data

We started our analysis with a hydrology dataset composed of time-series of daily mean flow from over 3800 gauging stations distributed throughout France, acquired from the HYDRO database (<http://www.hydro.eaufrance.fr/>). We removed stations from the dataset that lacked quality-assured data as defined by the HYDRO database managers, and for which the flow records were coded as modified due to the presence of reservoirs, diversions or significant abstractions in the upstream catchment. From the

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



remaining stations, we selected those for which flow data were available in the 35-yr period from 1975 to 2009. We searched these records for gaps longer than 20 days and removed the year of record in which these gaps occurred. This resulted in an average removal of 1.2 (std dev = 1.7) years per station. After these steps, 628 stations with 23–35 yr of record were retained (Fig. 2).

2.3 Flow variables and indices

We used the daily flow data for each gauging station to produce two annual time-series describing flow intermittence; the frequency of zero-flow periods (consecutive days of zero-flow) in each year of record (FREQ; yr^{-1}), and the total number of zero-flow days in each year of record (DUR; days). From these time-series we calculated two inter-annual indices for each station, the mean annual frequency of zero-flow periods (mFREQ; yr^{-1}), and the mean duration of zero-flow periods (mDUR; days). To determine whether zero flows may have been recorded at some gauging stations due to freezing, we also calculated mFREQ and mDUR separately for the winter season, December to February.

2.4 River network and environmental variables

Our spatial analysis used a digital representation of France's river network that was derived from a Digital Elevation Model (DEM) with 50-m resolution, made by Institut Géographique National (IGN) (Pella et al., 2012). The minimum catchment size used to define river segments in the DEM was 2.5 km^2 ; smaller catchments and the low-order segments in them were omitted from our analyses. The network comprised $\sim 115\,000$ uniquely identified segments (mean length 2.5 km) defined by upstream and downstream confluences with tributaries (Pella et al., 2012). We obtained available GIS layers comprising climatic, topographic and geological data for France and used these to define several environmental variables for each segment of the river network that have been shown previously to be associated with hydrological patterns (Snelder et

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



al., 2009) (see Table 1 for variable definitions). For each segment, the catchment area (the total area upstream; Area) and the unique subcatchment area (the area draining laterally into the segment) were delineated using the DEM.

Rainfall and air temperatures were measured by Météo-France at meteorological stations during the period 1961–1990 and were interpolated onto a 1-km grid (resolution 1 km) using the method of Benichou and Le Breton (1987). The climatic layers included average annual rainfall, mean monthly rainfall, average maximum daily temperature in the warmest month (July) and average minimum daily temperature in the coldest month (January). The temperature and rainfall data were then used to calculate values of the following variables for the catchment of each segment: average annual rainfall (Rain), the difference between average summer and winter rainfall divided by annual average rainfall (SumWinRain), average minimum January air temperature (T_{\min}), and average maximum July air temperature (T_{\max}) (Table 1). We derived two additional climate variables that described periods without rainfall: the catchment average of the mean annual number of days without rain (nDryDays), and the catchment average of the mean annual maximum duration of consecutive days without rain (dDry). The data used for these predictors were available only for stations, not for the entire network. Values of nDryDays and dDry were obtained from a rainfall time-series generated for the catchment of each gauging station derived from the high-resolution Safran atmospheric reanalysis over France (Quintana-Seguí et al., 2008) using methods described by Sauquet and Catalogne (2011).

Topographic data consisted of a slope grid that we derived from the DEM, and mapped river channels represented on the 1 : 250 000 scale BD Carthage[®] map obtained from IGN. We used the river-channel map to derive an estimate of drainage density that was independent of our DEM-based network. We calculated the catchment average values of slope and drainage density for each segment to define the variables Slope and Drain (Table 1). We also used the DEM to estimate the distance from each segment to the most distant point of its upstream catchment. This distance

HESSD

10, 1511–1551, 2013

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



divided by catchment area was defined as the variable Shape. Elongated catchments have high values of Shape, and round catchments have low values.

Geological data were derived from a 1:1 000 000-scale digital geological map of France obtained from Bureau des Recherches Géologiques et Minières (BRGM, 1996).

5 The map defined 22 categories and comprised ~ 18 000 individually categorized polygons with a mean area of 40 km². The map was used to develop two GIS layers describing physical hardness (i.e. resistance to erosion) and permeability. For each layer, each of the 22 geological categories was assigned an ordinal value corresponding to relative hardness and permeability (Table 2). For detailed methods (see Snelder et al.,
10 2008). We computed the catchment-average of the ordinal values representing physical hardness (Hard) and permeability (Perm) of geological categories. We also computed the proportions of each catchment occupied by the broad geological categories chalk (Chalk), limestone (Lime) and alluvium (Alluv) (Table 2). We also derived the average segment subcatchment values of the ordinal values representing hardness and permeability (segHard and segPerm) and the geological categories chalk (segChalk, segLime and segAlluv) to address the possibility that local geological conditions affect flow intermittence.

2.5 Flow-regime and intermittence classifications

We classified the 628 gauging stations retained for analysis into a flow-regime classification that separated intermittent and perennial stations, and a classification of the intermittent stations based on the frequency and duration of zero flow. For the flow-regime classification, stations with values of mFreq and mDur greater than zero were assigned to the intermittent class, and the remainder to the perennial class. In the intermittence classification, the intermittent stations from the previous classification were grouped into three subclasses that corresponded to subdivisions of a scatter plot of mFREQ versus mDUR. We grouped the 123 gauging stations on intermittent segments into three approximately equal-sized intermittence subclasses based on nominal thresholds for mFREQ and mDUR of 0.5 and 5 respectively. Subclass 1 was

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



characterized by low-frequency ($mFREQ < 0.5$), short-duration ($mDUR < 5$) zero-flow periods. Subclass 2 was characterized by low-frequency ($mFREQ < 0.5$), long-duration ($mDUR > 5$) zero-flow periods. Subclass 3 was characterized by high-frequency ($mFREQ > 0.5$), long-duration ($mDUR > 5$) zero-flow periods.

5 2.6 Spatial synchronization of intermittence patterns

To better understand the spatial scale of observed intermittence patterns, we evaluated the degree of spatial synchronization of the variables $FREQ$ and DUR at the gauging stations on intermittent segments using the Mantel statistic r (Mantel, 1967). The Mantel statistic is the Pearson correlation coefficient between two matrices of dissimilarities and is used to quantify spatio-temporal clustering (i.e. spatially organized synchronization). Our first dissimilarity matrix described the dissimilarity in annual intermittence patterns between pairs of stations. Our second dissimilarity matrix defined the geographic (Euclidian) distance between pairs of stations. The significance of the statistic is established by permutation (Legendre and Legendre, 1998).

Dissimilarities in the two annual times-series describing flow intermittence ($FREQ$ and DUR) between pairs of stations were calculated as $(1-\rho)$, where ρ is the rank (Spearman) correlation of the time series. Therefore, stations were compared on the basis of the relative frequency and duration of zero flow in each year, rather than the absolute magnitudes of the events. The calculation of dissimilarities was complicated by missing data for some years in the analysis period (1975–2009) due to gaps or because the station record had not commenced or had ended. Thus, we calculated the correlation between each pair of stations for the years in which data were available at both stations.

The assumption of linearity is implicit in Mantel tests and consequently, the tests will not detect spatial structures if there are non-linear relationships between space and synchronous behavior. For example, gauging stations in close proximity may have similar temporal behavior, but the behavior of widely separated pairs of stations may be unrelated. We used a Mantel correlogram to determine the level of synchrony among

HESSD

10, 1511–1551, 2013

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



the gauging stations at different spatial scales (Goslee and Urban, 2007). The Mantel correlogram used Mantel tests to determine the correlations between geographic distance and the indices FREQ and DUR for subsets of stations belonging to several distance classes. The distance classes subdivided the log-transformed distances between stations into nine equi-distant categories. Because of the multiple comparisons made by the Mantel correlogram, Bonferroni corrections were applied before interpreting the significance of the correlations (Goslee and Urban, 2007).

2.7 Statistical modeling of classifications

We used statistical classification models to relate the flow-regime and intermittence classifications to the environmental variables. First, we fitted a model that discriminated the 628 gauging stations on the basis of flow-regime class (perennial and intermittent) using the environmental variables as predictors. The statistical model was used to assess the degree to which intermittence was related to different environmental variables, and to make predictions of the flow-regime class for all segments (gauged and ungauged) in the river network. Second, we fitted a model that used the environmental variables to discriminate intermittence subclasses of the gauging stations classed as intermittent. This second model was used to assess the degree to which different temporal patterns in flow intermittence were related to different environmental variables.

We used Random Forest (RF) models to relate the flow-regime and intermittence classifications to the environmental variables. Recent studies have shown that RF models predict spatial patterns in river characteristics better than more conventional methods such as linear regression (Booker and Snelder, 2012; Snelder et al., 2012). For a detailed description of RF models (see Breiman, 2001 and Cutler et al., 2007). Briefly, an RF model comprises an ensemble of individual Classification and Regression Trees (CART, Breiman et al., 1984) that can be used in a classification mode to model the probability that each case belongs to some set of categories (here flow-regime and intermittence classes). In a classification context, CART partitions observations into groups that minimise the misclassification rate based on a series of binary

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



rules (splits) constructed from the predictor variables (here the environmental variables). CART models have two desirable features for modelling complex relationships: they are free from distributional assumptions, and they automatically fit non-linear relationships and high order interactions. However, CART models have two limitations, they do not produce an optimal tree structure and they are sensitive to small changes in input data (Hastie et al., 2001).

The limitations in CART models can be reduced by using RF models (Breiman, 2001). A final prediction of the probability that each case belongs to each category is based on the average of all the individual predictions obtained from the ensemble of trees (the forest). An important feature of RF models is that each tree is grown with a bootstrap sample of the training data. In addition, at each node only small, random samples of the predictors are used to define the split. The introduction of these random components, combined with averaging individual predictions over an ensemble of trees, increases the prediction accuracy of RF models while retaining the desirable features of CART.

RF models produce a limiting value of the generalization error (Breiman, 2001). As the number of trees (k) increases, the generalization error always converges to the minimum. Thus, RF models cannot be over-fitted (Cutler et al., 2007). The number of trees needs to be set sufficiently high to ensure that convergence occurs and this number depends on the number of variables that are used at each split. Model performance can be optimized by altering the number of trees and variables that are used at each split. However, we used the recommended default values (the square root of the total number of predictors) and a large number of trees (500; Cutler et al., 2007).

The structure of RF models can be examined using importance measures and partial dependence plots. Importance measures indicate the contribution of the predictors to model accuracy and are calculated from the degradation in model performance (i.e. the increase in misclassification rate) when a predictor is randomly permuted (Breiman, 2001). Importance represents the increase in the misclassification rate that could be expected for new cases (i.e. cases not used to fit the model) if the predictor was excluded

HESSD

10, 1511–1551, 2013

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



from the model (Breiman, 2001). Partial dependence plots show the marginal contribution of a predictor to the response (i.e. the response as a function of the predictor when the other predictors are held at their mean value; Friedman and Meulman, 2003). These plots are not a perfect representation of the effects of each predictor, particularly if there are interactions or predictors are strongly correlated, but they provide useful information for interpretation (Friedman and Meulman, 2003).

Predictors in an RF model may have positive importance values, even when their removal from the model does not cause a significant reduction in model performance. We used the procedure of Svetnik et al. (2004) to reduce the RF models to the most parsimonious set of predictors. The procedure uses a cross-validation process that recursively removes the least-important predictors from the model and tests if the reduced model has significantly lower prediction performance than the full model. We used the “1 standard error rule” (Breiman et al., 1984) to select the reduced model with the highest prediction performance that was not different, within the error generated from the cross-validation process, from the model with the best performance. The reduced models were considered to be the most parsimonious and we used these to interpret the relationships between predictors and the classifications.

We used a leave-one-out cross-validation procedure to estimate the performance of the models and to optimize the probability threshold for the flow-regime model. In the cross-validation step, we fitted RF models to as many subsets of the data as there were gauging stations. For each subset, we withheld one gauging station in turn from the training data used to fit the RF models. We then used the RF model to independently predict the probability of the withheld station belonging to each of the classes represented by the classification. The cross-validated probabilities were then converted into predictions of class membership for each station based on a chosen probability threshold.

The performance of a classification model is sensitive to the probability threshold that is applied (Freeman and Moisen, 2008). We used receiver operating curves (ROCs) to provide a method of evaluating the performance of the flow-regime model that was

HESSD

10, 1511–1551, 2013

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



independent of the threshold. ROC plots show the true positive rate (sensitivity) against the false positive rate (1-specificity) as the threshold varies from 0 to 1 (Hanley and McNeil, 1982). Good models have high true positive rates and relatively small false positive rates and, therefore, have curves that rise steeply at the origin, and level off near the maximum value of 1. The ROC plot for a poor model lies near the diagonal, where the true positive rate equals the false positive rate for all thresholds. The area under the ROC plot (AUC) is a measure of overall model performance that is independent of the threshold, with good models having an AUC near 1, while a poor models will have an AUC near 0.5 (Hanley and McNeil, 1982).

We used the cross-validation predictions for the flow-regime classification to derive ROC statistics to select the best threshold for assigning gauging stations to the perennial or intermittent class. There are several criteria that can be used to define the best threshold (Freeman and Moisen, 2008) including minimizing the misclassification rate and that maximising Cohen's kappa (Cohen, 1960). Kappa measures the agreement between two classifications, each of which classify N items into C mutually exclusive categories. We chose Kappa because it adjusts for chance agreement and is therefore a more robust measure than misclassification rate when observed occurrence is low (Freeman and Moisen, 2008). Kappa takes a value between 0 (no agreement) and 1 (complete agreement).

We also used the results of the cross-validation predictions and Kappa to characterise the performance of the intermittence classification model. We did not optimise the probability threshold for the intermittence classification and simply assigned stations to the intermittence sub-class with the highest probability.

HESSD

10, 1511–1551, 2013

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



3 Results

3.1 Zero-flows and indices

Of the 628 gauging stations, 123 had at least one zero-flow period in the period of record. No gauging stations had zero-flow periods exclusively in winter, indicating that freezing was not the sole cause of intermittence at any station. Zero-flows at the gauging stations used for our analysis were most frequent and longest in duration in 1976, followed by 1989–1991 and 2003 (Fig. 3). However, the year in which the highest zero-flow frequencies or longest zero-flow durations occurred at each station was variable. At least one station had its highest zero-flow frequency or longest zero-flow duration in almost every year (i.e. at least one value of 100% occurs in most years in Fig. 3).

The index mFREQ varied between 0 and 7.5 yr^{-1} across all 628 gauging stations. For stations on intermittent segments, mFREQ ranged from 0.03 to 7.5 with a mean of 0.6 and a median of 0.3, indicating that most intermittent segments had low zero-flow frequencies. For stations on intermittent segments, mDUR ranged from 1 to 128 days with a mean of 15 and a median of 7.3.

The two flow-intermittence indices mFREQ and mDUR were weakly, positively correlated ($r = 0.19$, $p = 0.03$; Fig. 4). The intermittence subclasses 1, 2 and 3 comprised 41, 40 and 42 stations respectively. Three stations that fell just outside the ranges of mFREQ and mDUR that defined the subclasses and were assigned to the closest subclasses (lower right quadrant in Fig. 4). Gauging stations on intermittent river segments occurred across France; the highest densities were located near the southern and western coasts (Fig. 2). There was no clear geographic pattern evident in the spatial distribution of the three flow intermittence types (Fig. 2).

3.2 Spatial synchronization

The correlation coefficients (Mantel r) between dissimilarities corresponding to FREQ and DUR values for gauging stations on intermittent segments, and their spatial

HESSD

10, 1511–1551, 2013

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



separation were 0.1 and 0.14 respectively ($p < 0.0001$; 10 000 permutations), indicating weak spatial synchronization. Mantel correlograms for both FREQ and DUR indicated that correlations were weak at all spatial scales (Fig. 5). The largest correlations were negative and occurred between stations with the largest separation (i.e. mean distances of 700 km).

3.3 Discrimination of intermittent and perennial gauging stations

The RF model that related the flow-regime classification of 628 gauging stations to environmental variables had a cross-validated AUC of 0.77. The performance of the RF model as measured by misclassification rate and Cohen's kappa was sensitive to the probability value used as the threshold for assigning stations to either class (Fig. 6). The minimum misclassification rate was 14.8% and occurred at a probability threshold of 0.49. The maximum Cohen's kappa was 0.47 and occurred at a threshold of 0.35 (Fig. 6).

The eight environmental variables retained by the reduced model, in order of importance, were T_{\min} , Rain, T_{\max} , Slope, WinSumRain, Shape, Area and the geological variable Perm (Fig. 7). The partial dependence plots indicated that the probability that a segment is intermittent increased with increasing T_{\max} , T_{\min} , and Shape, decreased with increasing Rain, Area, and Slope, and had a U-shaped response to Winsumrain and Perm (Fig. 7). The importance measures for the environmental variables retained by the reduced model varied from 0.6 to 1.1% (Fig. 7), indicating that the influence of each variable on model performance was similar, compared to the overall misclassification rate of 16%. The environmental variables nDryDays and dDry and the sub-catchment geology variables (subHard, subPerm, subChalk, subLime, subAlluv) did not contribute significantly to model performance.

When a probability threshold for the flow-regime model that maximised Cohen's kappa (0.35) was used, 39% of river segments represented by our river network were predicted to be intermittent (Fig. 8). When these predictions were aggregated into hydro-ecoregions, the proportion of intermittent segments in each region predicted

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



by the model had a high level of accuracy. The regression of observed proportion of intermittent segments versus predicted proportion of intermittent segments had an r^2 value of 0.73 (Fig. 9a).

The mapped predictions of the flow-regime model and the aggregation of these predictions into hydro-ecoregions highlighted a broad gradient in the probabilities of flow intermittence that corresponded to large-scale climate patterns (Figs. 8 and 9b). Regions with the highest proportion of intermittent segments tended to those with the lowest annual rainfall (Rain) < 800 mm, the highest winter temperature (T_{\min}) $> 5^\circ\text{C}$ (Fig. 9b) and the highest summer temperatures (T_{\max}) $> 20^\circ\text{C}$. Regions with high probabilities of flow intermittence were located along the Mediterranean and central Atlantic coasts, in the Midi-Pyrénées region (Fig. 8). In contrast, regions with higher annual rainfall (Rain) > 1100 mm, lower summer air temperature (T_{\max}) $< 16^\circ\text{C}$ and lower winter air temperatures (T_{\min}) $< 1^\circ\text{C}$ had high probabilities of perennial flow (Fig. 8b). There were additional, smaller regions in which river segments had high probabilities of flow intermittence. These regions were characterized by low values of hardness (Hard) and low slopes (Slope). There was also a general tendency for segments along large river mainstems to have low probabilities of belonging to the intermittent class (Fig. 8). Mainstem segments had large values of catchment area (Area) and generally had low values of Shape, both of which were associated with low probabilities of flow intermittence (Fig. 8).

3.4 Discrimination of flow-intermittence patterns

The RF model of the intermittence classification (classification of intermittent segments into three subclasses), had independent misclassification rates of 51, 47 and 31 % for classes 1, 2 and 3, respectively, and an overall misclassification rate of 46 %. The value of kappa for the comparison of predicted and actual membership of gauged river segments to the three intermittence types was 0.32.

The six significant environmental variables in the reduced intermittence classification model, in order of importance, were Area, Shape, T_{\min} , T_{\max} , Rain, and Slope

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



(Fig. 10). The importance values for the six environmental variables retained in the reduced model ranged from 1 to 2 %, indicating that the influence of each variable on model performance was similar, compared to the overall misclassification rate of 46 %. The environmental variables nDryDays and dDry and the subcatchment geology variables (subHard, subPerm, subChalk, subLime, subAlluv) did not contribute significantly to model performance.

The partial plots indicated that the intermittence subclasses were discriminated based on differences in catchment characteristics (Fig. 10). Subclass 1 (low-frequency, short-duration zero-flow periods) had highest probability of occurrence in large catchments with relatively high rainfall (Rain) and slope (Slope), cool summers (T_{\max}), cold winters (T_{\min}) and low Shape (i.e. rounded catchments). In Subclass 2 (low-frequency, long-duration zero-flow periods), the relationships with catchment characteristics were similar to those in Subclass 1, except the probability of occurrence decreased with increasing rainfall (Rain), and increasing Slope. Subclass 3 (high-frequency, long-duration zero-flow periods) had highest probability of occurrence in small catchments (Area) with low rainfall (Rain), high summer and winter air temperatures (T_{\max} and T_{\min}), and high Shape (i.e. elongated catchments).

4 Discussion

In this study we used statistical models to identify relationships between flow intermittence and catchment characteristics. To our knowledge, there have been no comparable studies that attempted to classify types of flow intermittence and predict its spatial distribution. Random forest models of the flow-regime and intermittence classifications had cross-validated kappa values of 0.47 and 0.32, respectively. Based on the guidelines of Fleiss (1981), these kappa values indicate that the performance of the flow-regime classification model was fair, and that of the intermittence classification model was poor. However, both classification models identified associations between

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



intermittence and environmental variables and provided some useful insights into the occurrence of flow intermittence.

The RF model of the flow-regime classification indicated that there were significant associations between intermittence and catchment rainfall, air temperature and catchment size, shape, slope and geology (Figs. 7 and 9). Regions with high probabilities of intermittent river segments are those with low annual rainfall, warm air temperatures, and steep, small, elongated catchments. The probability of intermittent river segments was also higher in catchments with either high or low rainfall in winter compared to summer, and high or low permeability. The environmental variables retained in the RF model of the flow-regime classification had similar importance values, i.e. they made comparable contributions to model performance. The similarity in importance values suggests that intermittence is caused by multiple physical factors, each of which has a moderate influence.

The RF model of the intermittence classification identified associations between environmental variables and different combinations of zero-flow frequency and duration (Fig. 10). Based on its relationship with the environmental variables, intermittence Subclass 3 appears to represent headwater streams in warm and dry locations that have frequent zero-flow periods. In contrast, Subclasses 1 and 2 appear to represent larger, cooler, and wetter catchments that have infrequent zero-flow periods (on average less than once per year ($mFREQ > 0.5$, Fig. 4). River segments in Subclasses 1 and 2 differed in that they had short- and long-duration zero-flow periods ($mDur > 5$ days and $mDur < 5$ days, Fig. 4), respectively. The most obvious environmental difference between Subclasses 1 and 2 was that sites in Subclass 1 had higher probability of occurrence in steeper catchments with higher annual rainfall (Slope and Rain, Fig. 10). It is likely that the low frequencies of zero-flow periods in Subclasses 1 and 2 are attributable to sustained base flows for most of each year, which are in turn associated with large catchment areas and relatively wet, cool climates. We suggest that zero-flow periods in these two sub-classes occur when the water table drops below the channel elevation and surface flow ceases. The difference in zero-flow duration

HESSD

10, 1511–1551, 2013

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



between Subclasses 1 and 2 may be associated with differences in storage, release and recharge of groundwater. Sites in Subclass 1 are likely to have small groundwater storage volumes due to steep slopes and small catchment areas. These volumes are rapidly depleted, but they are also recharged rapidly by rainfall, which keeps the duration of zero-flow periods short. The reverse applies to sites in Subclass 2, which generally had lower slopes and lower annual rainfall, producing zero-flow periods of long duration.

Our classification models indicated that the probability of intermittence in France is only broadly associated with large-scale climate patterns (Figs. 8 and 9b). Flow intermittence is at least partly controlled by processes acting at smaller-scales than climate, such as local groundwater-table fluctuations and seepage through permeable channels (Fleckenstein et al., 2006; Larned et al., 2010a,b). The proximate cause of flow intermittence in many river segments is water table fluctuation relative to river channel elevation; flow occurs when the water table intersects the channel, and ceases when the water table drops below the channel (Konrad, 2006; Larned et al., 2010a; von Schiller et al., 2011). In areas where channels are permanently perched above the water table, intermittence is controlled by run-off from upstream and by transmission losses (Morin et al., 2009; Sharma and Murthy, 1994). Interactions between these local processes and the climatic processes that generate runoff and rainfall-recharge determine where and when flow intermittence occurs. Two results lend support to the suggestion that intermittence, and in particular, the timing of zero-flow periods, are only partly influenced by large-scale climatic processes. First, the plots of annual zero-flow behavior (Fig. 3) and low values of Mantel statistics (Fig. 5) indicate that there was weak spatial synchronization in zero-flow frequency and duration. Although there was a general tendency for the frequency and duration of zero-flow periods to peak in the driest years of the time series (1976, 1989/1990, 2003), many sites did not follow this temporal pattern (Fig. 3). Second, the variables nDryDays and dDry represent rainfall patterns that we expected to be more relevant to the duration and frequency of zero flows than mean

HESSD

10, 1511–1551, 2013

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



annual rainfall, which was included in the RF models. However, these variables did not improve the classification models.

In contrast to the present study, a flow-regime classification of France reported by Snelder et al. (2009) discriminated river network segments on the basis of a variety of hydrological indices including the frequency of zero flow and others that described the frequency, timing and duration of high and low flows, and the frequency of changes in flow. A predictive model of this classification performed well and was based on similar statistical methods, and used the same river network and environmental variables used in the present study. We suggest that most of the flow-regime components described by the classification of Snelder et al. (2009) result predominantly from variation in large-scale processes such as rainfall, evaporation, catchment storage and runoff. Variation in these processes over large spatial scales was reflected in relatively strong discrimination of flow-regime classes using the environmental variables.

The contrasting performance of the models in the present study with the model of the flow-regime classification of Snelder et al. (2009) indicates that different suites of environmental variables are needed to model intermittent flows and whole flow-regimes (in which zero-flow frequency is only one of many components). Aquifer structure and riverbed permeability are probably among the factors that influence the small-scale processes that determine intermittence and were represented in our models by variables such as catchment slope, shape and permeability. These variables had similar levels of importance in the RF models as the variables representing broad-scale climate, which supports our contention that processes at a range of scales are involved. However, it is likely that, although these variables acted as surrogates for local processes, they were too broad-scale and insufficiently representative of the actual causes of intermittency to achieve good predictive performance. The inclusion of predictor variables that described subcatchment (i.e. local) geology (Hard, Perm, Chalk, Lime, Alluv) was intended to provide better surrogates for small-scaled processes; however, these did not improve model performance. It is likely that the 1 : 1 000 000 scale of the geological map used to define the geological predictor variables was too coarse to discriminate

HESSD

10, 1511–1551, 2013

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



geological or hydrogeological variation at segment-scales. If flow intermittence is related to groundwater dynamics, spatial data corresponding to aquifer structure, riverbed permeability and other small-scale factors are needed to model intermittent flows accurately, but these data are rarely available. Until that data scarcity is alleviated, predictive models of regional or national patterns in flow intermittence will have limited accuracy.

Our predictions of the abundance and distribution of intermittent rivers in France will have several potential practical applications. First, the Water Framework Directive requires the ecological status of all surface and ground waters to be assessed (Chave, 2001), and this includes intermittent rivers. However, partly because they have been considered “atypical” and rare in France, intermittent rivers have been often ignored by water managers. Our predictions may increase awareness of the prevalence of intermittent rivers among scientists and water managers by showing they are abundant and occur in most regions of France and are not restricted to the dry Mediterranean region.

Our flow-regime classification predictions indicate that the gauging station network under-represents intermittent river segments in France (i.e. 19.6% of gauges were classified as intermittent whereas 39% of segments in the river network are predicted to be intermittent). More accurate predictions of the abundance and distribution intermittent segments could be achieved by reconfiguring the gauging network. Alternative methods of monitoring flow intermittence, such as the use of electrical resistance arrays (Jaeger and Olden, 2012) or citizen-observation networks (Turner and Richter, 2011) could also increase the representation of intermittence in future studies at less effort than is required to operate permanent gauging stations.

The predicted proportion of intermittent segments by our study is particular to the spatial resolution of the analysis (i.e. the minimum catchment area to define a segment was 2.5 km² and the minimum catchment area of our gauging stations was 3 km²). Based on the relationship between catchment area and intermittent segments (Fig. 7), a finer-resolution analysis would likely result in higher estimates of the proportion of intermittent segments. Other studies have also concluded that lower order streams are

HESSD

10, 1511–1551, 2013

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



more likely to be intermittent and represent a large proportion of river networks by length (Meyer et al., 2007).

Although predictions of intermittence were not accurate at the segment scale, when aggregated by HER they accurately predicted the proportion of intermittent segments at regional scales (Fig. 9). Regional estimates of the proportion of intermittent segments provide important information for several management applications including environmental flow-setting (Hughes, 2005), estimating assimilative capacity for contaminants (Tsagarakis et al., 2004) and designing bio-monitoring programs that are representative of the full range of river environments (Steward et al., 2011).

Finally, predictive models describing intermittence can be used to provide preliminary estimates of the how climate change could change the frequency of intermittence (e.g. Benito et al., 2011). Our results suggest that the probability of intermittence in France would typically increase by $\sim 2\%$ with each 1°C rise in summer air temperature (T_{max}) and by 3% for each 100-mm reduction in mean annual rainfall (Rain) (Fig. 7).

Acknowledgements. Ton Snelder was supported by Marie Curie Incoming International Fellowship within the 6th European Community Framework Programme and by a joint Irstea-Onema research project on Low Flow Hydrology). We thank André Chandèsris for assistance with hydrological data and the French Water Agency Rhône-Méditerranée-Corse for financial support. We acknowledge the use of data supplied by the French database Banque Hydro. Ton Snelder and Scott Larned were funded by the New Zealand Foundation for Research, Science and Technology, Environmental Flows Programme (C01X0308).

References

- Acuña, V., Muñoz, I., Giorgi, A., Omella, M., Sabater, F., and Sabater, S.: Drought and post-drought recovery cycles in an intermittent Mediterranean stream: structural and functional aspects, *J. Am. Water Resour. Assoc.*, 24, 919–933, 2005.
- Angel, R., Asaf, L., Ronen, Z., and Nejidat, A.: Nitrogen transformations and diversity of ammonia-oxidizing bacteria in a desert ephemeral stream receiving untreated wastewater, *Microb. Ecol.*, 59, 46–58, 2010.

HESSD

10, 1511–1551, 2013

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Regionalization of patterns of flow intermittence

T. H. Snelder et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[◀](#)

[▶](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



- Arcott, D. B., Larned, S., Scarsbrook, M. R., and Lambert, P.: Aquatic invertebrate community structure along an intermittence gradient: Selwyn River, New Zealand, *J. Am. Water Resour. Assoc.*, 29, 530–545, 2010.
- Benichou, P. and Le Breton, O.: Prise en compte de la topographie pour la cartographie des champs pluviométriques statistiques (Use of topography on mapping of statistical rainfall fields), *La Météorologie*, 7, 23–34, 1987.
- Benitoa, G., Thorndycraft, V. R., Rico, M. T., Sánchez-Moya, Y., Sopeña, A., Botero, B. A., Machado, M. J., Davis, M., and Pérez-González, A.: Hydrological response of a dryland ephemeral river to southern African climatic variability during the last millennium, *Quaternary Res.*, 75, 471–482, 2011.
- Bent, G. C. and Steeves, P. A.: A revised logistic regression equation and an automated procedure for mapping the probability of a stream flowing perennially in Massachusetts, US Department of the Interior, US Geological Survey, Massachusetts, 2006.
- Booker, D. J. and Snelder, T. H.: Comparing methods for estimating flow duration curves at ungauged sites, *J. Hydrol.*, 434–435, 78–94, 2012.
- Breiman, L.: Random Forests, *Mach. Learn.*, 45, 5–32, 2001.
- Breiman, L., Friedman, J. H., Olshen, R., and Stone, C. J.: *Classification and Regression Trees*, Wadsworth, Belmont, California, 1984.
- B.R.G.M.: Carte géologique de France au 1/1.000.000ème, Bureau de Recherches Géologiques et Minières, Paris, France, 1996.
- Brooks, R. T. and Colburn, E. A.: Extent and channel morphology of unmapped headwater stream segments of the Quabbin Watershed, Massachusetts, *J. Am. Water Resour. Assoc.*, 47, 158–168, 2011.
- Chave, P.: *The EU Water Framework Directive: An Introduction*, IWA Publishing, London, 2001.
- Cohen, J.: A coefficient of agreement for nominal scales, *Educ. Psychol. Meas.*, 20, 37–46, 1960.
- Corti, R. and Datry, T.: Invertebrate and sestonic matter in an advancing wetted front travelling down a dry riverbed (Albarine, France), *Freshwater Sci.*, 31, 1187–1201, 2012.
- Corti, R., Datry, T., Drummond, L., and Larned, S.: Leaf litter decomposition along the advancing-retreating front of a temporary river, *Aquat. Sci.*, 73, 537–550, 2011.
- Crocker, K. M., Young, A. R., Zaidman, M. D., and Rees, H. G.: Flow duration curve estimation in ephemeral catchments in Portugal, *Hydrolog. Sci. J.*, 48, 427–439, 2003.

HESSD

10, 1511–1551, 2013

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Cutler, D. R., Edwards, J. T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., and Lawler, J. J.: Random forests for classification in ecology, *Ecology*, 88, 2783–2792, 2007.
- Datry, T.: Benthic and hyporheic invertebrate assemblages along a flow intermittence gradient: effects of duration of dry events, *Freshwater Biol.*, 57, 563–574, 2012.
- 5 Datry, T., Arscott, D., and Sabater, S.: Recent perspectives on temporary river ecology, *Aquat. Sci.*, 73, 453–457, 2011.
- Datry, T., Corti, R., and Philippe, M.: Spatial and temporal aquatic–terrestrial transitions in the temporary Albarine River, France: responses of invertebrates to experimental rewetting, *Freshwater Biol.*, 57, 716–727, 2012.
- 10 Davey, A. J. H. and Kelly, D. J.: Fish community responses to drying disturbances in an intermittent stream: a landscape perspective, *Freshwater Biol.*, 52, 1719–1733, 2007.
- Dieter, D., von Schiller, D., Garcia-Roger, E. M., Sánchez-Montoya, M. M., Gómez, R., Mora-Gómez, J., Sangiorgio, F., Gelbrecht, J., and Tockner, K.: Preconditioning effects of intermittent stream flow on leaf litter decomposition, *Aquat. Sci.*, 73, 599–609, 2011.
- 15 Elmore, A. J. and Kaushal, S. S.: Disappearing headwaters: patterns of stream burial due to urbanization, *Front. Ecol. Environ.*, 6, 308–312, 2008.
- Fleckenstein, J. H., Niswonger, R. G., and Fogg, G. E.: River-aquifer interactions, geologic heterogeneity, and low-flow management, *Ground Water*, 44, 837–852, 2006.
- Fleiss, J. L.: *Statistical methods for rates and proportions*, John Wiley, New York, 223 pp., 1981.
- 20 Freeman, E. A. and Moisen, G. G.: A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa, *Ecol. Model.*, 217, 48–58, 2008.
- Friedman, J. H. and Meulman, J. J.: Multiple additive regression trees with application in epidemiology, *Stat. Med.*, 22, 1365–1381, 2003.
- 25 Fritz, K. M., Johnson, B. R., and Walters, D. M.: Physical indicators of hydrologic permanence in forested headwater streams, *J. Am. Water Resour. Assoc.*, 27, 690–704, 2008.
- Gómez, R., Hurtado, I., Suárez, M., and Vidal-Abarca, M.: Ramblas in south-east Spain: threatened and valuable ecosystems, *Aquat. Conserv. Mar. Freshw. Ecosyst.*, 15, 387–402, 2005.
- Goslee, S. C. and Urban, D. L.: The ecodist package for dissimilarity-based analysis of ecological data, *J. Stat. Softw.*, 22, 1–19, 2007.
- 30 Hanley, J. A. and McNeil, B. J.: The meaning and use of the area under a receiver operating characteristic (ROC) curve, *Radiology*, 143, 29–36, 1982.

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Hansen, W. F.: Identifying stream types and management implications, *Forest Ecol. Manage.*, 143, 39–46, 2001.
- Hastie, T., Tibshirani, R., and Friedman, J. H.: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer-Verlag, New York, 2001.
- 5 Heine, R. A., Lant, C. L., and Sengupta, R. R.: Development and comparison of approaches for automated mapping of stream channel networks, *Ann. Assoc. Am. Geogr.*, 94, 477–490, 2004.
- Houston, J.: Variability of precipitation in the Atacama Desert: its causes and hydrological impact, *Int. J. Climatol.*, 26, 2181–2198, 2006.
- 10 Hughes, D. A.: Hydrological issues associated with the determination of environmental water requirements of ephemeral rivers, *River Res. Appl.*, 21, 899–908, 2005.
- Jacobson, P. and Jacobson, K.: Hydrologic controls of physical and ecological processes in Namib Desert ephemeral rivers: Implications for conservation and management, *J. Arid Environ.*, doi:10.1016/j.jaridenv.2012.01.010, in press, 2012.
- 15 Jaeger, K. L. and Olden, J. D.: Electrical resistance sensor arrays as a means to quantify longitudinal connectivity of rivers, *River Res. Appl.*, 28, 1843–1852, 2012.
- Ji, X., Kang, E., Chen, R., Zhao, W., Zhang, Z., and Jin, B.: The impact of the development of water resources on environment in arid inland river basins of Hexi region, Northwestern China, *Environ. Geol.*, 50, 793–801, 2006.
- 20 Katz, G. L., Denslow, M. W., and Stromberg, J. C.: The Goldilocks effect: intermittent streams sustain more plant species than those with perennial or ephemeral flow, *Freshwater Biol.*, 57, 467–480, 2012.
- Kennard, M. J., Pusey, B. J., Olden, J. D., Mackay, S. J., Stein, J. L., and Marsh, N.: Classification of natural flow regimes in Australia to support environmental flow management, *Freshwater Biol.*, 55, 171–193, 2010.
- 25 Kikawada, T., Minakawa, N., Watanabe, M., and Okuda, T.: Factors inducing successful anhydrobiosis in the African chironomid *Polypedilum vanderplanki*: significance of the larval tubular nest, *Integr. Comp. Biol.*, 45, 710–714, 2005.
- Konrad, C.: Longitudinal hydraulic analysis of river-aquifer exchanges, *Water Resour. Res.*, 42, 8425, doi:10.1029/2005WR004197, 2006.
- 30 Lake, P.: Ecological effects of perturbation by drought in flowing waters, *Freshwater Biol.*, 48, 1161–1172, 2003.

HESSD

10, 1511–1551, 2013

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

- Larned, S. T., Arscott, D. B., Schmidt, J., and Diettrich, J. C.: A Framework for Analyzing Longitudinal and Temporal Variation in River Flow and Developing Flow-Ecology Relationships, *J. Am. Water Resour. Assoc.*, 46, 541–553, 2010a.
- Larned, S. T., Detry, T., Arscott, D. B., and Tockner, K.: Emerging concepts in temporary-river ecology, *Freshwater Biol.*, 55, 717–738, 2010b.
- Larned, S. T., Schmidt, J., Detry, T., Konrad, C. P., Dumas, J. K., and Diettrich, J. C.: Longitudinal river ecohydrology: flow variation down the lengths of alluvial rivers, *Ecohydrology*, 4, 532–548, 2011.
- Legendre, P. and Legendre, L.: *Numerical ecology*, Elsevier, Amsterdam, The Netherlands, 1998.
- Leibowitz, S. G., Wigington Jr., P. J., Rains, M. C., and Downing, D. M.: Non-navigable streams and adjacent wetlands: addressing science needs following the Supreme Court's Rapanos decision, *Front. Ecol. Environ.*, 6, 364–371, 2008.
- Leopold, L. B.: *A View of the River*, Harvard Univ Press, California, 1994.
- Mantel, N.: The detection of disease clustering and a generalized regression approach, *Cancer Res.*, 27, 209–220, 1967.
- Meirovich, L., Ben-Zvi, A., Shentsis, I., and Yanovich, E.: Frequency and magnitude of runoff events in the and Negev of Israel, *J. Hydrol.*, 207, 204–219, 1998.
- Meyer, J. L., Strayer, D. L., Wallace, J. B., Eggert, S. L., Helfman, G. S., and Leonard, N. E.: The contribution of headwater streams to biodiversity in river networks, *J. Am. Water Resour. Assoc.*, 43, 86–103, 2007.
- Morin, E., Grodek, T., Dahan, O., Benito, G., Kulls, C., Jacoby, Y., Langenhove, G. V., Seely, M., and Enzel, Y.: Flood routing and alluvial aquifer recharge along the ephemeral arid Kuiseb River, Namibia, *J. Hydrol.*, 368, 262–275, 2009.
- Nadeau, T. L. and Rains, M. C.: Hydrological connectivity between headwater streams and downstream waters: how science can inform policy, *J. Am. Water Resour. Assoc.*, 43, 118–133, 2007.
- Olden, J. D., Kennard, M. J., and Pusey, B. J.: A framework for hydrologic classification with a review of methodologies and applications in ecohydrology, *Ecohydrology*, 5, 503–518, 2011.
- Pella, H., Lejot, J., Lamouroux, N., and Snelder, T.: Le réseau hydrographique théorique (RHT) français et ses attributs environnementaux, *Géomorphologie*, 3, 317–336, 2012.



Regionalization of patterns of flow intermittence

T. H. Snelder et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

- Perry, S., Euverman, R., Wang, T., Loong, A., Chew, S., Ip, Y., and Gilmour, K.: Control of breathing in African lungfish (*Protopterus dolloi*): a comparison of aquatic and cocooned (terrestrialized) animals, *Resp. Physiol. Neurobiol.*, 160, 8–17, 2008.
- 5 Quintana-Segui, P., Le Moigne, P., Durand, Y., Martin, E., Habets, F., Baillon, M., Canellas, C., Franchisteguy, L., and Morel, S.: Analysis of near surface atmospheric variables: Validation of the SAFRAN analysis over France, *J. Appl. Meteorol. Clim.*, 47, 92–107, 2008.
- Sauquet, E. and Catalogne, C.: Comparison of catchment grouping methods for flow duration curve estimation at ungauged sites in France, *Hydrol. Earth Syst. Sci.*, 15, 2421–2435, doi:10.5194/hess-15-2421-2011, 2011.
- 10 Sayer, M. D. J.: Adaptations of amphibious fish for surviving life out of water, *Fish Fish.*, 6, 186–211, 2005.
- Sharma, K. and Murthy, J.: Estimating transmission losses in an arid region – a realistic approach, *J. Arid Environ.*, 27, 107–112, 1994.
- Snelder, T. H., Pella, H., Wasson, J., and Lamouroux, N.: Definition procedures have little effect on performance of environmental classifications of streams and rivers, *Environ. Manage.*, 15 42, 771–788, 2008.
- Snelder, T. H., Lamouroux, N., Leathwick, J. R., Pella, H., Sauquet, E., and Shankar, U.: Predictive mapping of natural flow regimes of France, *J. Hydrol.*, 373, 57–67, 2009.
- Snelder, T. H., Ortiz, J. B., Booker, D., Lamouroux, N., Pella, H., and Shankar, U.: Can bottom-up procedures improve the performance of stream classifications?, *Aquat. Sci.*, 74, 45–59, 20 2012.
- Steward, A. L., Marshall, J. C., Sheldon, F., Harch, B., Choy, S., Bunn, S. E., and Tockner, K.: Terrestrial invertebrates of dry river beds are not simply subsets of riparian assemblages, *Aquat. Sci.*, 73, 551–566, 2011.
- 25 Svetnik, V., Liaw, A., Tong, C., and Wang, T.: Application of Breiman's Random Forest to Modeling Structure-Activity Relationships of Pharmaceutical Molecules, 5th International Workshop on Multiple Classifier Systems, Springer, Cagliari, Italy, 334–343, 2004.
- Tsagarakis, K. P., Dialynas, G., and Angelakis, A.: Water resources management in Crete (Greece) including water recycling and reuse and proposed quality criteria, *Agr. Water Manage.*, 66, 35–47, 2004.
- 30 Turner, D. S. and Richter, H. E.: Wet/dry mapping: using citizen scientists to monitor the extent of perennial surface flow in dryland regions, *Environ. Manage.*, 47, 497–505, 2011.

von Schiller, D., Acuña, V., Graeber, D., Martí, E., Ribot, M., Sabater, S., Timoner, X., and Tockner, K.: Contraction, fragmentation and expansion dynamics determine nutrient availability in a Mediterranean forest stream, *Aquat. Sci.*, 73, 485–497, 2011.

5 Wasson, J. G., Chandesris, A., Pella, H., and Blanc, L.: Typology and reference conditions for surface water bodies in France – the hydro-ecoregion approach, *Proceeding of Typology and Ecological Classification of Lakes and Rivers*, Nordic Council of Ministers, 25–26, 2002.

HESSD

10, 1511–1551, 2013

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Regionalization of
patterns of flow
intermittence

T. H. Snelder et al.

Table 1. Environmental variables used as predictors in statistical classification models. Values are means and ranges (in brackets) for the gauging stations (i.e. the data used to fit the models) and the network (i.e. the entire domain to which the models were extrapolated). Variables were the catchment average for each segment of the river network. The geological descriptors; Hard, Perm, Allu, Chalk and Lime were also derived for the sub-catchment of each segment.

Variable	Description	Stations	Network
T_{\min}	Average minimum air temperature of catchment in January ($^{\circ}\text{C}$).	-1, (-12; 5)	-0.5, (-15; 7)
T_{\max}	Average maximum air temperature of catchment in July ($^{\circ}\text{C}$).	24, (11; 31)	24, (9; 32)
Rain	Average annual rainfall (mm).	1030, (606; 2060)	944, (507; 2536)
SumWinRain	Difference of average winter (December, January and February) and summer (June, July and August) rainfall, divided by the average annual rainfall.	0.1, (-0.1; 0.2)	0.1, (-0.2; 0.3)
Area	Area of catchment (km^2).	1010, (3; 108 893)	684, (3; 117 500)
Slope	Mean hill slope of the catchment (m km^{-1}).	15, (1; 75)	11, (0; 103)
Drain	Mean drainage density of the catchment (km km^{-2}).	1, (0; 2.7)	0.9, (0; 3.3)
Hard	Catchment averaged value of hardness.	3.9, (1; 5)	3.5, (0; 5)
Perm	Catchment averaged value of permeability.	2, (1; 4)	2.3, (0; 4)
Allu	Proportion of catchment with alluvium surface geology.	0.06 (0; 0.97)	0.09 (0; 1)
Chalk	Proportion of catchment with alluvium surface geology.	0.06 (0; 0.99)	0.11 (0; 1)
Lime	Proportion of catchment with alluvium surface geology.	0.03 (0; 0.99)	0.05 (0; 1)
Shape	Distance to head of catchment divided by catchment area ($\text{km km}^{-2} \times 10^3$).	0.18 (0.01; 0.8)	0.3 (0; 3.5)
nDryDays	Mean annual number of days without rain.	229 (194; 301)	NA
dDry	Mean annual maximum duration without rain	21.6 (17; 41)	NA

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



HESSD

10, 1511–1551, 2013

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[◀](#)

[▶](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



Table 2. Geological categories represented on the 1 : 1 000 000 scale geological map of France (BRGM, 1996). Ordinal values of physical hardness range from 1 (soft) to 5 (hard); ordinal values of permeability range from 1 (impermeable) to 4 (permeable). The fourth column shows the geological categories that were also included independently as environmental variables in our analysis (as proportion of the catchment in category). The sum of three categories was used to define alluvium (Allu).

Geological category	Physical hardness	Permeability	Catchment geological category
Fluvial alluvium	1	3	Allu
Quaternary alluvium	1	3	Allu
Clay and sand	1	1	
Limestone	5	4	Lime
Chalk	3	4	Chalk
Glacial deposit	1	1	
Sedimentary flysch	3	2	
Marls	3	1	
Marls with evaporates	3	1	
Molasse	2	3	
Basaltic rock	5	1	
Igneous rock	5	1	
Calcareous detrial rock	4	2	
Crystalline detrial rocks	5	1	
Noncalcareous detrial rock	5	3	
Metamorphic rocks	5	1	
Volcanic rocks	5	2	
Sand	1	3	Allu
Carbonaceous schist	3	2	
Metamorphic schist	3	2	
Sedimentary schist	3	2	
Stratified calcareous rocks	3	2	

HESSD

10, 1511–1551, 2013

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

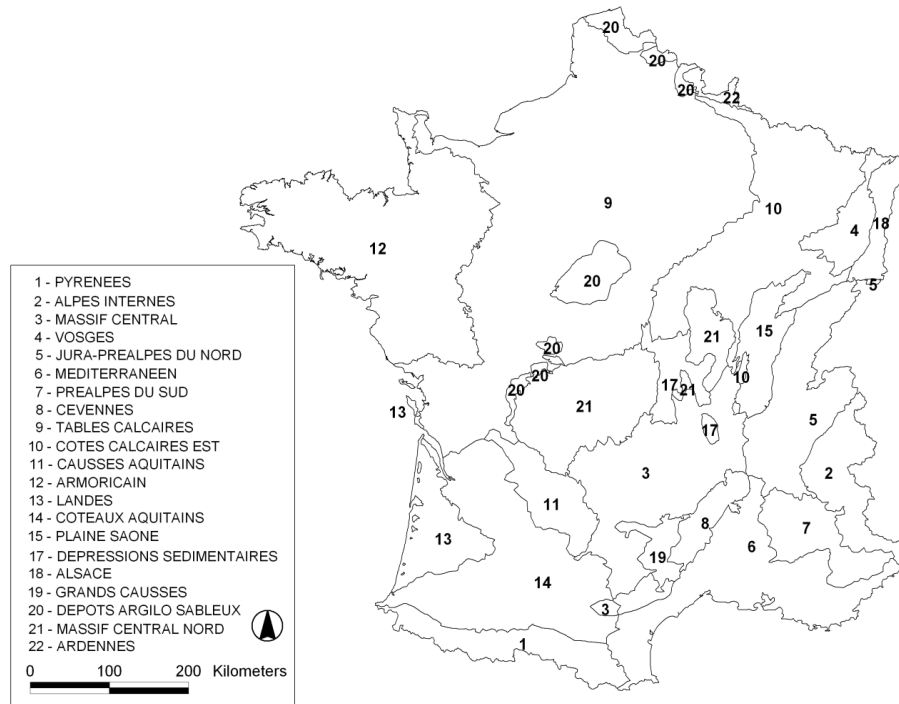


Fig. 1. France showing the HER boundaries.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

Regionalization of
patterns of flow
intermittence

T. H. Snelder et al.

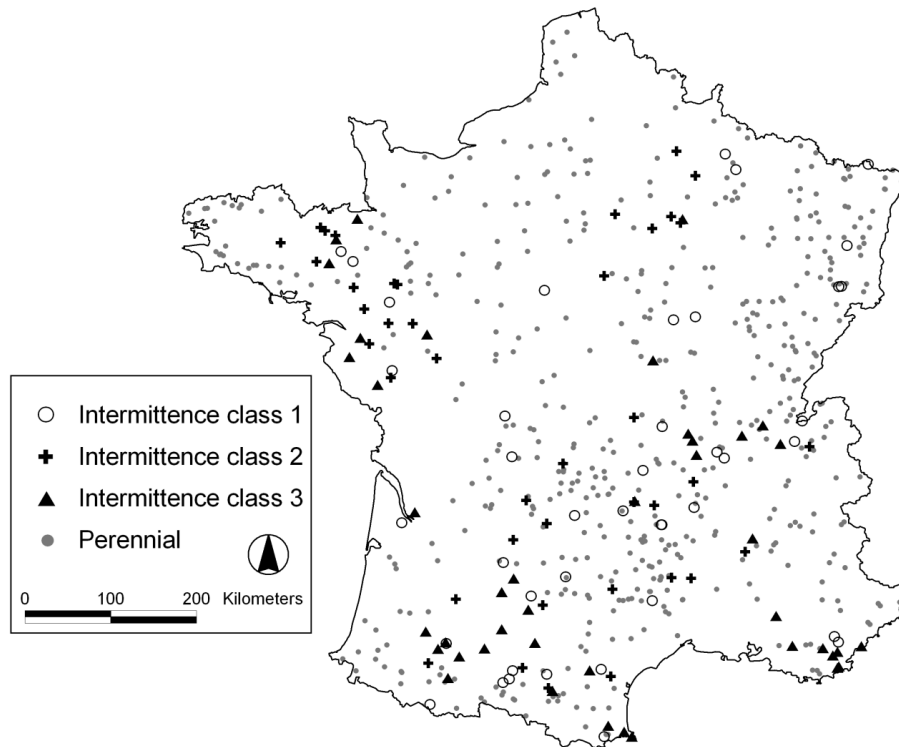


Fig. 2. Map of the location of the gauging stations used in this study. Stations on perennial segments are indicated by closed circles. Stations on intermittent segments are classified into three intermittence classes (see text for class descriptions).

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

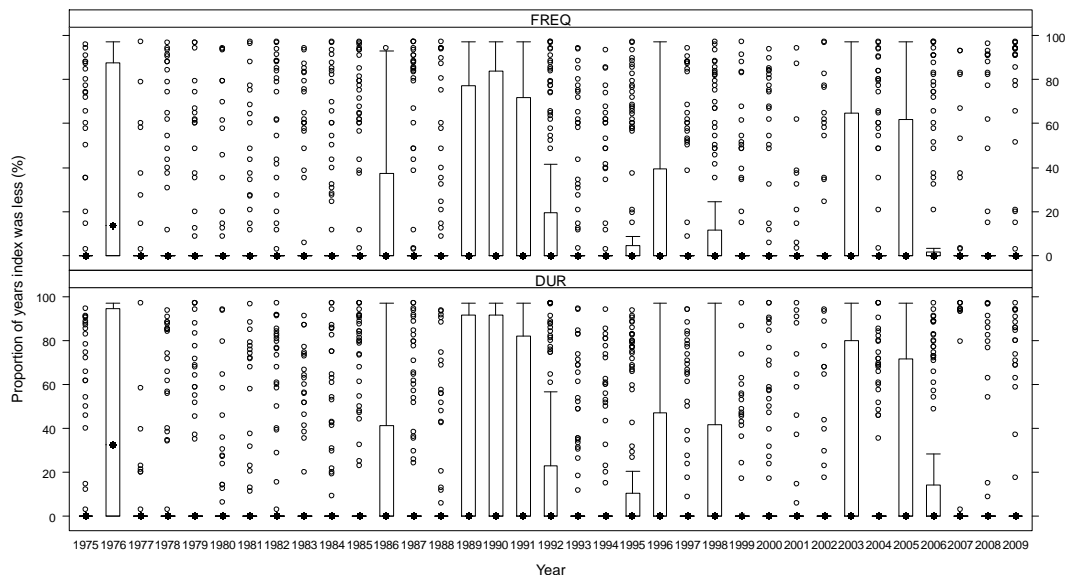


Fig. 3. Box plots of the annual variables (FREQ and DUR) for all years. The plots show data for only the gauging stations on intermittent segments. The variables have been standardized within stations by expressing them as the frequency that the index was less. Plotted values therefore indicate the severity of the zero-flow events in each year relative to the extremes observed at each station. The box contains the inter-quartile range, the dot shows the median value, whiskers indicate 1.5 times interquartile range and the circles indicate outliers.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Regionalization of
patterns of flow
intermittence

T. H. Snelder et al.

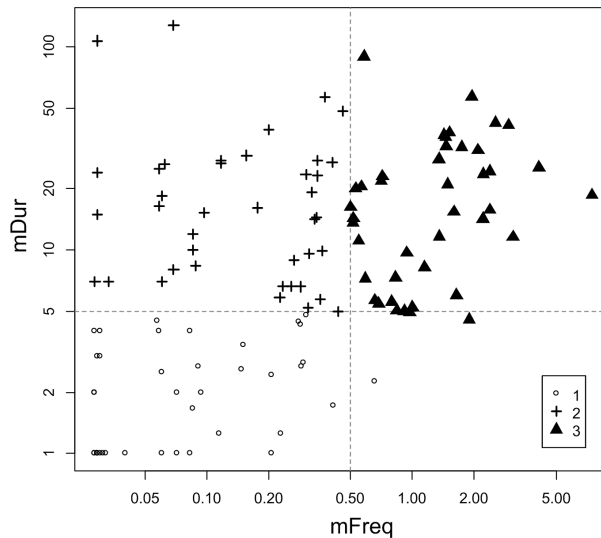


Fig. 4. Intermittence index values for gauging stations on intermittent segments. Three intermittence subclasses are indicated by different symbols. Subclasses were defined using threshold values of mFREQ and mDUR (dashed lines).

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[⏪](#)[⏩](#)[⏴](#)[⏵](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

Regionalization of
patterns of flow
intermittence

T. H. Snelder et al.

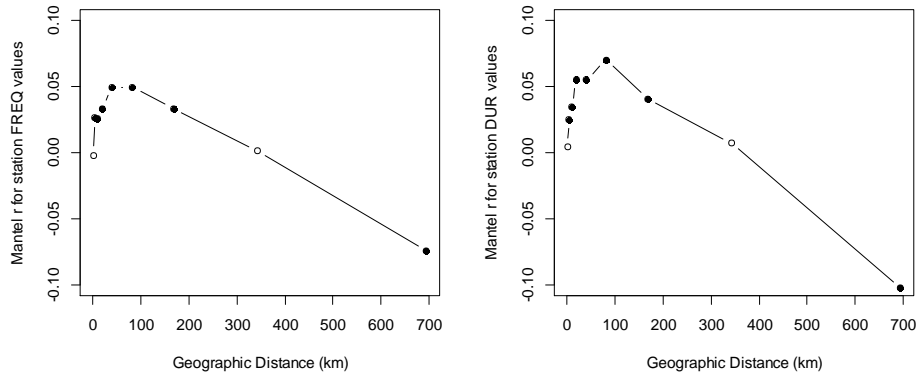


Fig. 5. Mantel correlograms of FREQ and DUR for gauging stations on intermittent segments in nine distance classes. The plots indicate the Mantel r -values for each distance class plotted against the mean separation of gauging stations in the class. Solid dots indicate significant Mantel r -values ($p < 0.05$) after Bonferroni correction.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



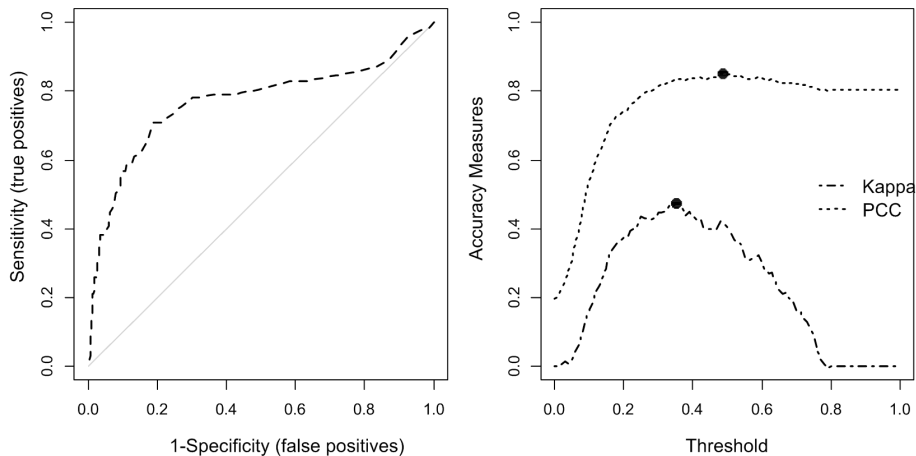


Fig. 6. ROC plot (left panel) and threshold plot (right panel) for the flow-regime classification.

Regionalization of patterns of flow intermittence

T. H. Snelder et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Regionalization of
patterns of flow
intermittence

T. H. Snelder et al.

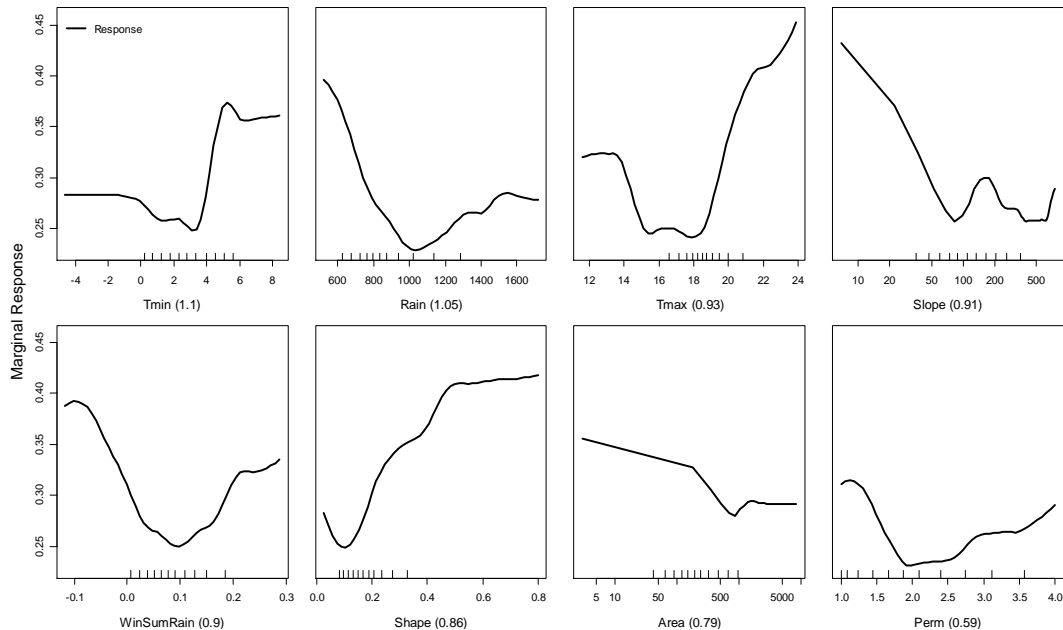


Fig. 7. Partial dependence of the probability of intermittence for each of significant environmental variables retained in the reduced model. The variables are shown in order of importance from the top-right to the bottom left. The plots show the marginal contribution to probability of flow intermittence (y-axis) as a function of the variables (i.e. the other contributing variables were held at their mean). The rug plots on the horizontal axes show deciles of the predictors. The value in brackets on the horizontal axis is the importance measure.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



**Regionalization of
patterns of flow
intermittence**

T. H. Snelder et al.

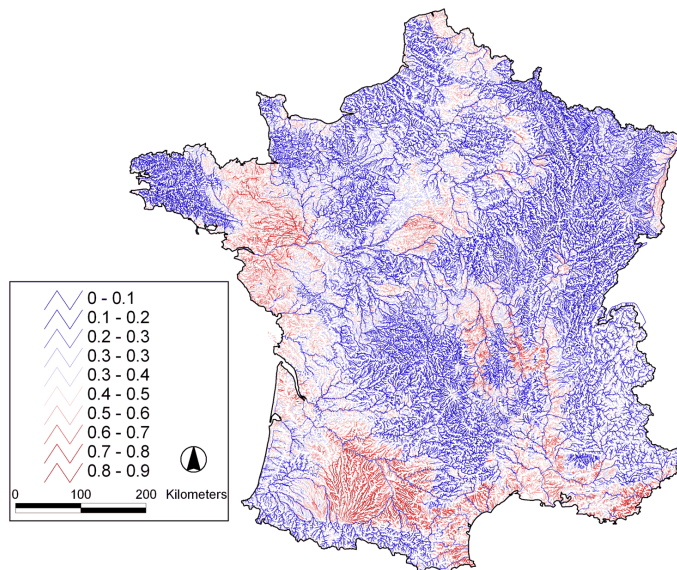


Fig. 8. Predictions of the probability of being intermittent made by the RF model of the flow-regime classification for the entire river network.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

Regionalization of
patterns of flow
intermittence

T. H. Snelder et al.

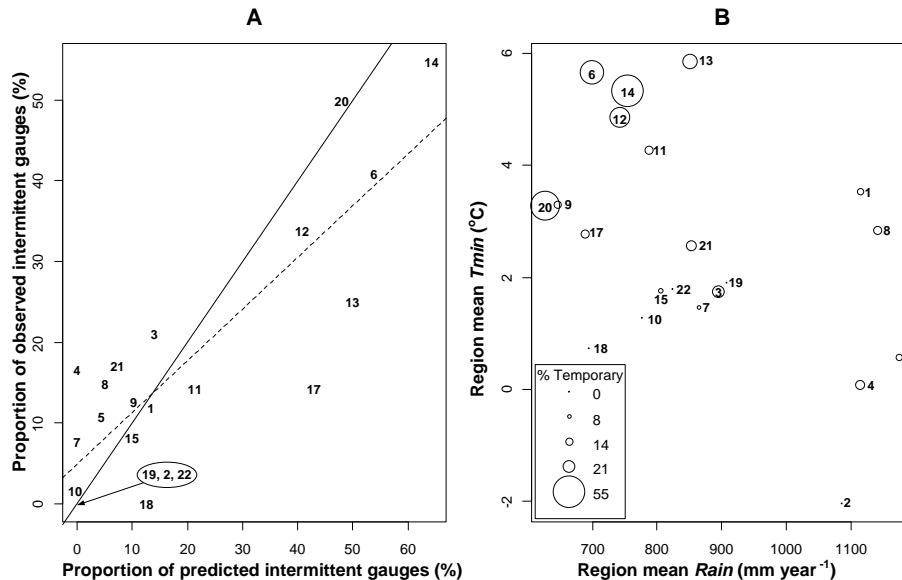


Fig. 9. (A) Observed and predicted proportions of gauging stations on intermittent segments in hydro-ecoregions. The dashed line is the regression of observed versus predicted gauges on intermittent segments. (B) Scatter plot of the mean values of the most important predictor variables (T_{min} and Rain) in each HER. The symbol sizes indicate the percentage of gauges on intermittent segments each region.

Regionalization of
patterns of flow
intermittence

T. H. Snelder et al.

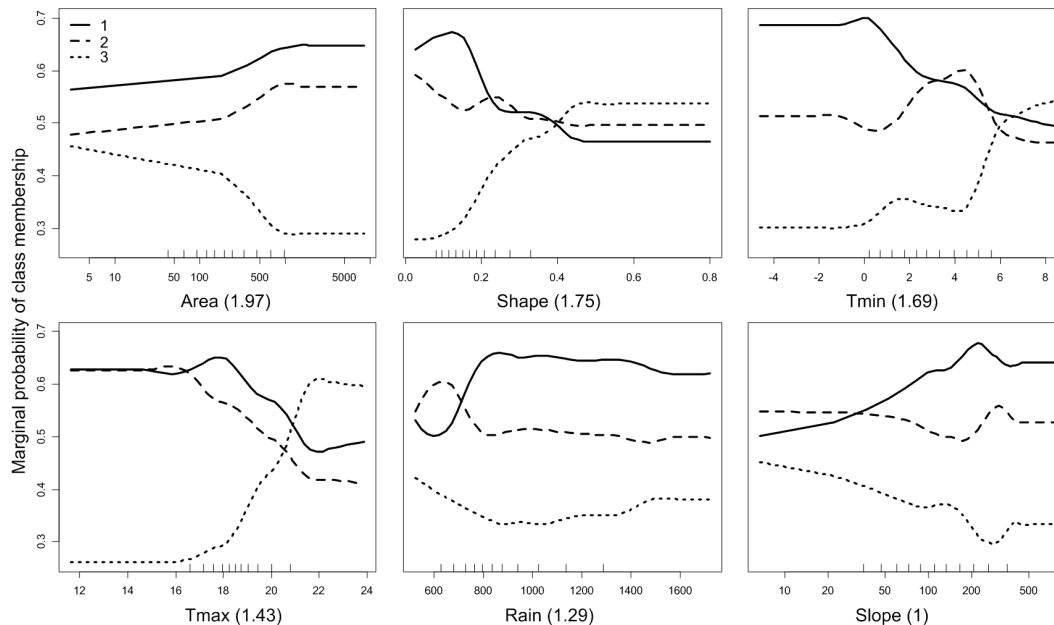


Fig. 10. Partial dependence plots for the three flow intermittence classes for the six environmental variables retained in the reduced model, in order of importance from the top-right to the bottom left. The plots show the marginal contribution to probability of class membership (y-axis) as a function of the variable (i.e. the other contributing variables were held at their mean). The rug plots on the horizontal axes show deciles of the predictors. The value in brackets on the horizontal axis is the importance measure.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

